

NO-A191 917

A PROBABILISTIC/POSSIBILISTIC APPROACH TO MODELING C3
SYSTEMS PART 2(U) NAVAL OCEAN SYSTEMS CENTER SAN DIEGO
CR I R GOODMAN NOV 87

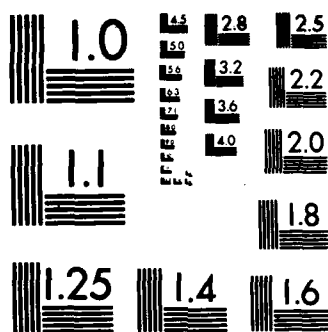
1/1

UNCLASSIFIED

F/G 25/3

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

A number of results involving this general form are presented, including: justification for use of expert-derived information for inference rules and other factors/tie-ins with plausibility measures; characterizations of formal language symbolisations and related data fusion results; and a new approach to data fusion evaluation through algebraic logic, developing a formal counterpart to conditional probabilities- "conditional objects" for consistent manipulation of disparate data.

Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



A PROBABILISTIC / POSSIBILISTIC APPROACH TO MODELING C^3 SYSTEMS: PART II

I.R. Goodman

Command & Control Department
Code 421

NAVAL OCEAN SYSTEMS CENTER
San Diego, California 92152

ABSTRACT

This paper continues the work begun in the last Proceedings (9th MIT/ONR Workshop on C^3 Systems). In that work, C^3 systems are considered as interacting networks of decision-making node complexes characterized by system or process variables. Internodal relations are modeled through nonlinear additive (in the general sense) regression relations; intranodal relations are made to follow a general SHOR (Sense-Hypothesize-Option-Response) paradigm. In turn, it is shown that a collection of ten types of relatively primitive implication or conditional relations PRIM between C^3 variables for enemy and friendly component systems determines all updated marginal node state distributions. (Distributions can be in the classical probabilistic sense or more generally in a multi-valued logical sense.) This leads to a C^3 decision game, where the loss function is some picked combination of measures of performance or effectiveness derived from node states and where each decision strategy corresponds to some choice of PRIM for each C^3 system.

In the present work, emphasis is placed upon model refinement. In particular, the intranodal relation representing data fusion is expanded and analyzed. This expansion is characterized by a weighted sum of products for the classical probability case and extended to a more general form for multi-valued logics. A number of results involving this general form are presented, including: justification for use of expert-derived information for inference rules and other factors/tie-ins with plausibility measures; characterizations of formal language symbolizations and related data fusion results; and a new approach to data fusion evaluation through algebraic logic, developing a formal counterpart to conditional probabilities—"conditional objects" for consistent manipulation of disparate data.

1. INTRODUCTION

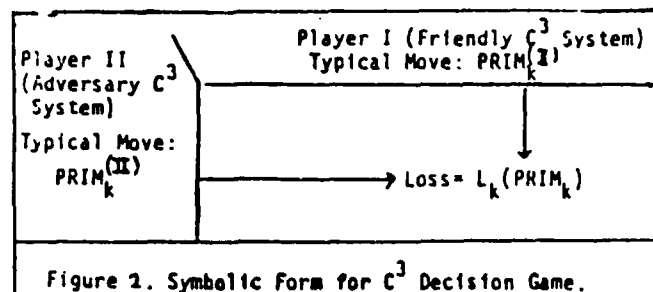
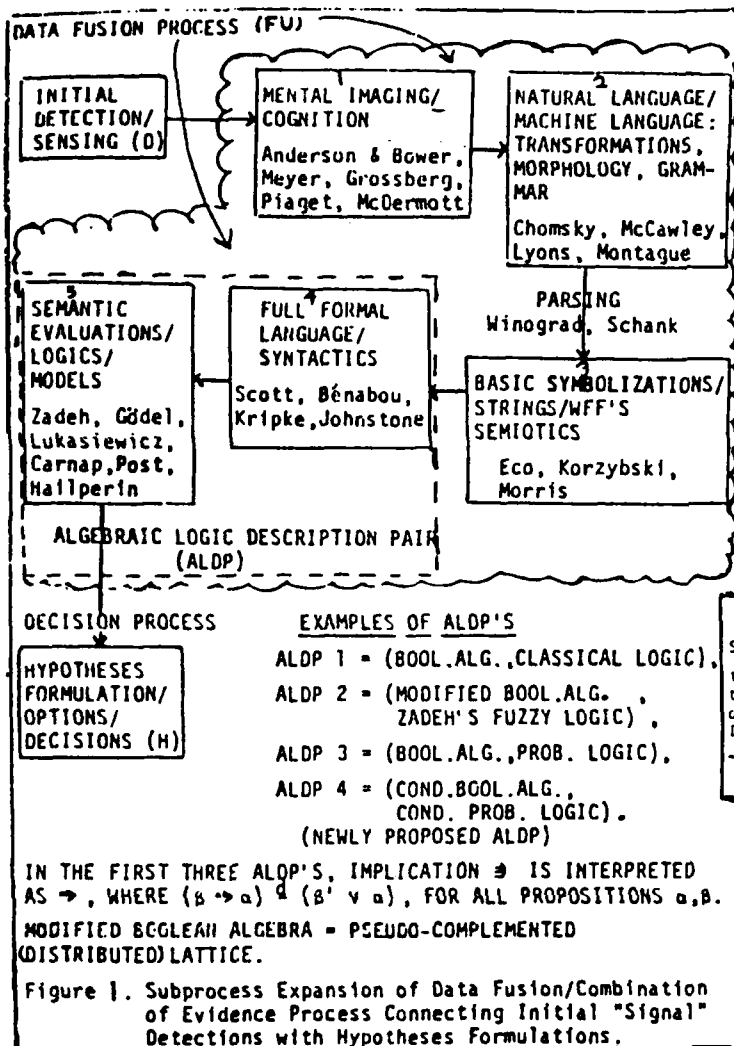
This paper, for the most part, is an abridgment of a much longer version [1].

For the past several years, throughout many fields of science and technology, researchers have been seeking unification and extension of past results in order to explain reality better and to be able to predict future developments. Recent events in theoretical physics involving "superstring" theory, an attempt at developing a Grand Unified Theory of the Universe, underscore this quest [2].

In a more modest way, this paper seeks to establish a theory unifying, coordinating, and extending the somewhat-appearing distinct concepts of data fusion, combination of evidence, and C^3 systems analysis. On the other hand, relatively little attention will be paid here to detailed computational techniques which are particular to certain types of common data fusion problems such as regression procedures for combining stochastic sensor information, or maximum likelihood or Bayesian procedures for putting together geolocation data arriving from different sources relative to a given target of interest. All of the above-mentioned techniques are essentially special cases of a much more general combination of evidence approach, on which this paper will concentrate.

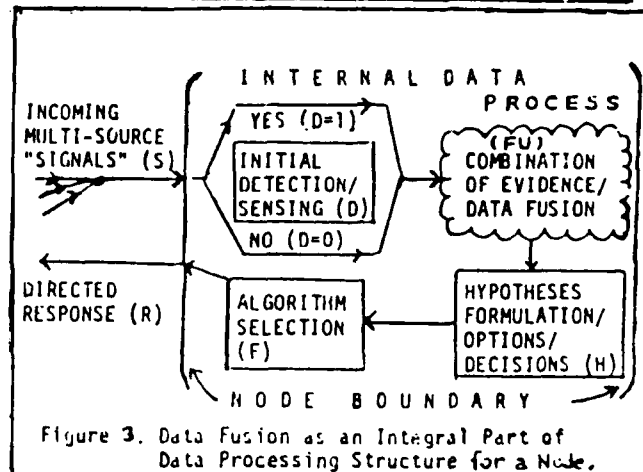
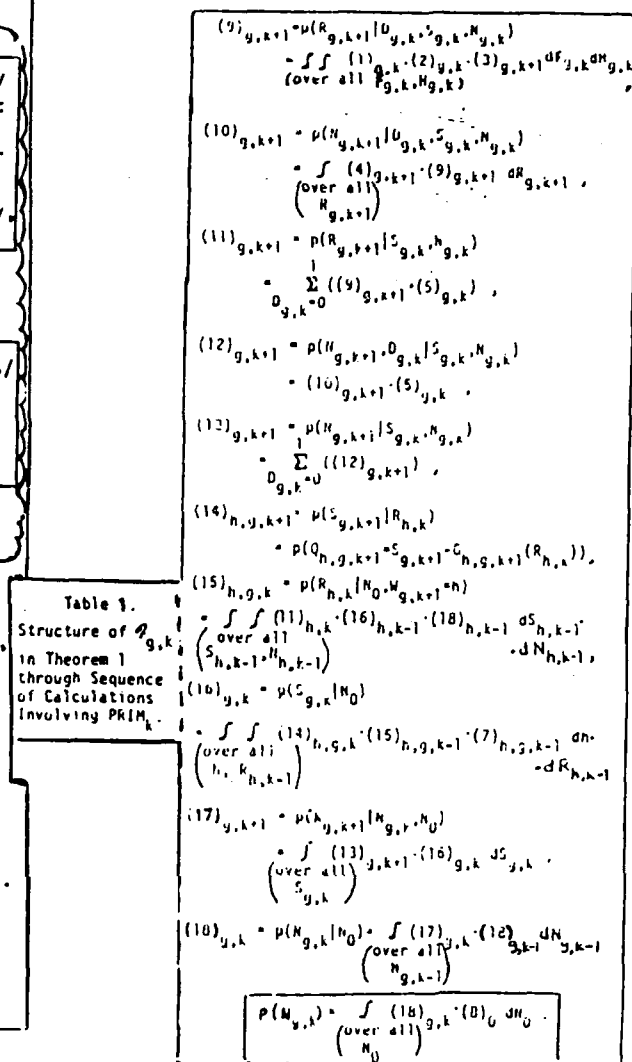
Previously, this author proposed a bottoms-up, microscopic, quantitative approach to general C^3 systems [4]. In that approach, a generic C^3 system is identified as a network of node complexes of decision-makers, human or automated, interfacing with each other in general. Each node receives "signals" which may be ordinary communication signals, either from friendly or hostile sources (possibly unaware), or which may be received weapon fire. In general, these "signals" are stacked vectors comprised of incoming data from several different nodes. In turn, each node—which may consist of a single decision-maker or some coalition of decision-makers and which may include passive type decision-makers, such as "followers"—then processes the data. This is followed by a response or action taken towards other nodes, friendly or hostile. ([1], Figure 1.) Associated with each node is the node state ([1], Figure 2.) describing the current state-of-affairs given in terms of a number of functions such as threat level, equations of motion, and supply level. In addition, there is an associated knowledge base reflecting the node's local knowledge of the other nodes (friendly or adversary). Also associated with each node is its internal "signal" processing design, as described in Figure 3. There, data fusion plays a central role in transmitting detected "signals" to hypotheses formulations, which in turn through algorithm selection leads to an output response to other nodes (again, these may be friendly or adversary).

Next, since we identify data fusion with the combining of evidence, all of the knowledge-based system techniques associated with the latter are available. In particular, this infers (see [5], Chapters 1, 2 and Figure 1, page 14) that a series of underlying processes are involved in data fusion. Basically, there are five such processes (including natural language in its broadest context) given in Figure 1.



2. DATA FUSION AS A QUANTITATIVE PART OF AN OVERALL C^3 SYSTEM AND DECISION GAME

So far, in this development toward a general theory for the fusion of data, only general qualitative descriptions have been given for the processes involved. However, as mentioned before, a quantitative model for generic C^3 systems has been established compatible with these qualitative formulations [4]. Inputs to the structure consist basically of ten sorts of known relative primitive relations $PRIM$ among the variables describing a C^3 system. These variables are: node(N); hypotheses selection (H); detection



(1) of incoming "signals" (S); algorithm selections (F); initial node responses (R), prior to environmental distortion (G) and additive noise (Q). To each variable is affixed subscripts (g,k) (or {h,g,k}) where $g(a,i)$ denotes the identification of a particular

node in question in terms of the C^3 system a (friendly or hostile) and node number l , while k represents a discrete time index t . Specifically, the relation breaks down into 5 intranodal (within nodes) relations, 2 internodal (between nodes) or regression relations, and 3 prior relations for each C^3 system. These relations are expressed in terms of conditional or unconditional probabilities, as they stand, but the results can be extended, with appropriate replacements, to a multivalued logic setting. (Again, see [4].) Then by making certain reasonable sufficiency assumptions among the variables and utilizing basic properties of conditional probabilities, it can be shown that each updated node state can be obtained explicitly in (probabilistic) terms of the other variables and node states through PRIM. Thus, we have:

Theorem 1. (See [4], Theorem 1.)

Suppose $PRIM_k$ and $N_{g,k}$ are as described above with $PRIM_k$ given in further details in Table 2.

Then, under certain reasonable sufficiency conditions [4],

$$p(N_{g,k}) = \Phi_{g,k}(PRIM_k), \quad (2.1)$$

where $\Phi_{g,k}$ is a computable functional involving a finite number of integrations and arithmetic operations upon the elements of $PRIM_k$ given in Table 1.

PRIM _k	INTEGRAL	(1) _{g,k} $\triangleq p(N_{g,k} D_{g,k}, S_{g,k})$, (3) _{g,k+1} $\triangleq p(R_{g,k+1} t_{g,k}, S_{g,k}, N_{g,k})$
		(2) _{g,k} $\triangleq p(F_{g,k} H_{g,k})$, (4) _{g,k+1} $\triangleq p(N_{g,k+1} R_{g,k+1}, H_{g,k})$
		(5) _{g,k} $\triangleq p(D_{g,k} S_{g,k}, N_{g,k})$
	PRIOR	(8) _{g,0} $\triangleq p(N_{g,0})$, (15) _{h,g,0} $\triangleq p(R_{h,0} W_{g,0}, h, N_{g,0})$
		(16) _{g,0} $\triangleq p(S_{g,0} N_{g,0})$
	INTER-NODAL	(6) _{h,g,k+1} $\triangleq p(Q_{h,g,k+1})$ with $G_{h,g,k+1}$
		(7) _{h,g,k+1} $\triangleq p(W_{h,g,k+1} h N_{g,0})$

The basic internodal analysis is developed via additive nonlinear regression relation
 $(S_{g,k+1}|W_{g,k+1}, (h,k)) = G_{h,g,k+1}(R_{h,k+1}) + G_{h,g,k+1}$
 where variable $W_{g,k+1}$ indicates original possible possible node source for "signal" at time k , given reception by another node at $k+1$.

Table 2.
Relative Primitive Relations for C^3 System a .
 $g(a, i)$, where i indicates node l .
 k indicates time index.

In turn, a simple two-person zero sum game can be established, called the C^3 decision game. Here, Player I corresponds to entire C^3 system $a=1$ (say, friendly) and Player II corresponds to entire C^3 system $a=2$ (say, adversary). In this game, a move by Player j corresponds to a choice (up to given constraints) of $PRIM_k^{(j)}$, $j=1, II$, and the resulting loss or utility due to any such joint move L_k is a function of the marginal updated node state distributions, according to Theorem 1 as

$$L_k(PRIM_k) = MOE_k(\{p(N_{g,k})\} \{all\} g)) \\ = MOE_k(\{\Phi_{g,k}(PRIM_k)\} \{all\} g)), \quad (2.2)$$

where MOE_k represents a single figure-of-merit, combining various measures of effectiveness (moe 's) or performance (mop 's) for the two C^3 systems. (Note, that although ideally the entire joint node state distribution of the two C^3 systems should be sought, in practice this is difficult to do, because of the great combinatoric computations involved.) Typical moe 's

that could be used include: averaged measure of importance $TH_{a,k}$; averaged measure of threat $\overline{TH}_{a,k}$; upper bound total entropy $ENT_{a,k}$; and averaged measure of performance $ACC_{a,k}$, all computable through $p(N_{g,k})$'s for C^3 system a , by use of Theorem 1. (See also [4], eqs. (59)-(62).) Then one could let

$$MOE_k = MOE_{1,k} - MOE_{2,k}, \quad (2.3)$$

where

$$MOE_{a,k} = \lambda_1 \cdot \overline{TH}_{a,k} + \lambda_2 \cdot \overline{TH}_{a,k} + \lambda_3 \cdot \overline{ENT}_{a,k} + \lambda_4 \cdot \overline{ACC}_{a,k}, \quad (2.4)$$

and the λ_i 's are some predetermined weightings.

Symbolically, the C^3 decision game appears as given in Figure 2.

Finally, one can then apply all the usual game-theoretic methods to this C^3 game, such as seeking Bayes decision functions for moves, least favorable strategies (all subject to practical constraints), minimax strategies, the game value, and various sensitivity measures. It is the long-range hope that such a game will be a useful decision-aid in planning command strategy. At present, a relatively simple implementation scheme is being carried out for testing the feasibility of such an approach to C^3 systems.

3. STRUCTURE FOR DATA FUSION: THE CLASSICAL PROBABILITY CASE

With the general C^3 system context for data fusion established in the previous sections, let us now return to the task of developing a general quantitative structure for data fusion. In light of the previous remarks (again, see Figure 3), fusion is a process intermediate with initial sensing and hypotheses formulations, within a C^3 node complex of decision-makers. In addition, the fusion process decomposes into natural subprocesses (see Figure 1). Thus, in essence, we wish to expand the first relative primitive intranodal relation appearing in Table 2:

$$P(FU) = p(H|D, S), \quad (3.1)$$

where for reasons of convenience from now on we suppress the denotational-time indices, unless necessary. As stated before, p need not necessarily refer to ordinary probability evaluation, but may represent other evaluations such as possibilities for Zadeh's Fuzzy Logic or for more general multivalued truth systems.

In determining the above evaluation, another variable Z is often present. Z represents the vector of auxiliary or "nuisance" characteristics or attributes which can be useful in connecting H , the variable representing possible hypotheses or decisions as to what unknown parameter value or situation or diagnosis is occurring, with input data S and detection state D . Thus for example, if we are physically in a bunker- a C^3 node- S may be observed loud noise, with $D=1$ (definitely detected), and H could have possible domain values say $\text{dom}(H) = (H_1, \dots, H_5)$ as given in Table 3

- H_1 = no change in previous situation
- H_2 = enemy is about to mount the promised big offense
- H_3 = enemy is just feeling us out
- H_4 = enemy wants to negotiate
- H_5 = none of the above situations hold

Table 3. Typical Set of Values for $\text{dom}(H)$.

Thus, $\text{dom}(H)$ could serve as a legitimate sample space, if conditional probability $p(H|D, S)$ could be obtained for all possible values of H in $\text{dom}(H)$, i.e. $(H|D, S)$ could be interpreted as a random variable over $\text{dom}(H)$. In this case, suppose also that Z is an auxiliary variable representing any of a likewise collection of disjoint exhaustive situations locally going on at the bunker. Here, let $\text{dom}(Z)$ be given as in Table below:

- Z_1 = nothing happening
- Z_2 = accidental explosion in compartment #1
- Z_3 = accidental explosion in compartment #2
- Z_4 = enemy shot missile at us and it either hit us or just missed
- Z_5 = none of the above situations hold

Table 4. Typical Set of Values for $\text{dom}(Z)$.

Thus, again by disjointness and exhaustion, it is reasonable to conclude that $\text{dom}(Z)$ could serve as a legitimate sample space and Z can be interpreted as a random variable. All of this leads to the evaluation of the conditional probabilities $p(Z|D, S)$, which together with the values for $P(H|D, S)$ can be used to obtain the standard "integrated-out" form for the posterior distribution of H as given below:

$$p(H=H_j|D\&S) = \sum_{i=1}^5 p(H_j \& Z_i | D\&S) \\ = \sum_{i=1}^5 p(Z_i | D\&S) \cdot p(H_j | Z_i \& D\&S), \quad (3.2)$$

using the standard chaining property of conditional probabilities and replacing the antecedent comma notation by conjunctions. One could reasonably interpret the evaluation in (3.2) as the probability value for the expression

$$\text{"If } D \text{ and } S, \text{ then } H_j \text{"} \quad (3.3)$$

through the probability values for the expressions

$$\text{"If } D \text{ and } S, \text{ then } Z_i \text{" and "If } Z_i \text{ and } D \text{ and } S, \text{ then } H_j \text{"} \quad (3.4)$$

Of course, one need not use the above evaluation exactly to obtain useful equivalent values. As it stands, $P(Z_i|D\&S)$ can be interpreted as an error or variability probability for attribute Z , while $p(H_j|Z_i \& D\&S)$ can be understood to mean the inference rule probability connecting Z and D and S with H . On the other hand, often the conditional data or regression probability $p(S|Z_i \& H_j)$ and the joint prior probability $p(Z_i \& H_j)$

are available, assuming here $D=1$, which by use of Bayes' theorem also yields $p(H=H_j|D\&S)$. One standard result is to assume the above probabilities are gaussian, which in the discrete problem here, must serve as very rough approximations- in addition, the sets $\text{dom}(H)$ and $\text{dom}(Z)$ are not easily ordered compatible with a real domain for gaussian random variables. Then, if the mean of the conditional data distribution is linear in the data S , $p(H_j \& Z_i | S)$ takes on a generalized weighted least squares form. (See, e.g. [6].) The final result, $p(H=H_j|S)$, as in (3.2), is then a mixture of the probabilities of such least squares estimators.

4. STRUCTURE FOR DATA FUSION: THE CLASSICAL PROBABILITY CASE MODIFIED

Retaining the same terminology as before, suppose now that H, Z, S are variables such that any of the corresponding "sample spaces" do not truly contain disjoint exhaustive events; in particular, the disjointness condition may be violated more often than exhaustiveness- which we will assume here is always satisfied. Then it follows that simple corresponding probability measures as in Section 4 cannot be immediately assigned. Nor should "brute-force" normalization procedures be employed, unless absolutely necessary. For example, consider H . Suppose in the above example in Section 3 (Table 3), the enemy could simultaneously mount the promised offense (H_2), yet also be feeling us out for peace (H_3), or, even additionally, wanting to negotiate (H_4). Thus, in that case, $\text{dom}(H) = (H_1, \dots, H_5)$, as it stands, is not a suitable sample space of disjoint elementary events. Indeed, the elementary events H_i are not so elementary, many of them, due to complex causes, being overlapping! Equivalently, H in its current form may not be a legitimate random variable. What to do?

In particular, consider the crucial expression Q for data fusion appearing as primitive intranodal relation (1) in Table 2, sans the probability evaluation, and in natural language form:

$$Q \text{ "If } D \& S, \text{ then } H \text{"}. \quad (4.1)$$

In symbolic form, where $\&$ represents "and", \vee represents "or", $()$ represents "not", \Rightarrow represents implication,

$$Q = (D \cdot S \Rightarrow H). \quad (4.2)$$

Theorem 2. (See [1], Theorem 4.)

Suppose a formal language of propositions satisfies constraints (a), (b), (c), (d). Suppose also that variables D, S, H, Z are to be interpreted as before in the general sense and are such that (i) and (ii) are satisfied*, where constraints (*) are given in [1], Section 6. Then:

$$Q = \bigvee_{Z_i \in \text{dom}(Z)} g(Z_i; D, S; H), \quad (4.3)$$

where for all Z_i in $\text{dom}(Z)$,

$$g(Z_i; D, S; H) \stackrel{\text{def}}{=} (D \cdot S \Rightarrow Z_i \cdot H) \\ = g(Z_i; D, S) \cdot h(H; Z_i; D, S), \quad (4.4)$$

where

$$g(Z_i; D, S) = (D \cdot S \Rightarrow Z_i) \quad (4.5)$$

can be interpreted as an attribute variability or error form and

$$h(H; Z_i; D, S) = (Z_i \cdot D \cdot S \ni H) \quad (4.6)$$

can be interpreted as an inference rule connecting Z_i and H .

Given variables D, S, H and auxiliary variable Z ;
Next, for convenience define for all i, j

$$\alpha_i \triangleq (D \cdot S \ni Z) ; \alpha_j \triangleq (D \cdot S \ni Z_j) ; \quad (4.7)$$

$$\beta_j \triangleq (Z \cdot D \cdot S \ni H) ; \beta_{ij} \triangleq (Z_i \cdot D \cdot S \ni H_j). \quad (4.8)$$

$$\text{dom}(\alpha) = \{\alpha_i | i \in I\} \quad \text{dom}(Z) = \{Z_i | i \in I\}, \quad (4.9)$$

$$\begin{aligned} \text{dom}(\beta) &= \{\beta_{ij} | i \in I, j \in J\} \quad \text{dom}(Z) \times \text{dom}(H) \\ &= \{(Z_i, H_j) | i \in I, j \in J\}. \end{aligned} \quad (4.10)$$

$$A_j \triangleq \{(\alpha_i, \beta_{ij}) | i \in I\} = \{(Z_i, Z_i, H_j) | i \in I\}. \quad (4.11)$$

Theorem 3. ([5], Chapter 5) (See [1], Theorem 5.)

Let $\text{poss} : \text{dom}(\cdot) \rightarrow [0, 1]$ be any function, perhaps representing the expert opinions of a panel, as human integrators of information, taking into account the complex and possible overlapping natures of the events in $\text{dom}(\cdot)$.

Then make the following semantic evaluation of Q preserving the formal structure in Theorem 2:

$$\begin{aligned} \text{poss}(Q = Q_j) &= \text{poss}(Q = (D \cdot S \ni H_j)) \\ &= \bigoplus_{i \in I} (\bigoplus_{j \in J} (\text{poss}_{\alpha}(\alpha_i), \text{poss}_{\beta}(\beta_{ij}))). \end{aligned} \quad (4.12)$$

Then:

$$\begin{aligned} \text{poss}(Q = Q_j) &= \text{poss}(A_j \cap (S_{\alpha} \times S_{\beta}) \neq \emptyset) \\ &= p(A_j \cap (S_{\alpha} \times S_{\beta}) \neq \emptyset) \\ &= \text{plaus}_{S_{\alpha} \times S_{\beta}}(A_j), \end{aligned} \quad (4.13)$$

where $\text{plaus}_{S_{\alpha} \times S_{\beta}}$ denotes the plausibility or upper

probability measure with respect to random subset $S_{\alpha} \times S_{\beta}$ of $\text{dom}(\alpha) \times \text{dom}(\beta)$.

Remarks.

For related results, see the multivalued logic and fuzzy set approach to correlation and tracking through the PACT algorithm [3]. For general background, see [5], Ch. 3, 4. Shafer [7] developed use of plausibility measures and other bijectively related functions, such as "belief" and "doubt" measures in modeling combination of evidence problems. However, Nguyen [8] has emphasized, via Choquet's Capacity Theorem, which characterizes such functions in terms of both their random set connections and their generalized Poincaré expansion forms, that such "measures" require full specification of the associated random (sub)sets.

5. STRUCTURE FOR DATA FUSION: THE GENERAL COMBINATION OF EVIDENCE CASE

Let us return to the formal language aspect of data fusion as given in Theorem 3. In general know-

ledge-based systems, such as medical diagnosis ones, consist of a collection of inference rules corresponding to $h(H; Z_i; D, S)$ linking either observed data, such as D, S or portions of intermediate variable Z with other portions of Z or with diagnoses directly, played by the role of variable H . Similar comments hold for the attribute variability term $g(Z_i; D, S)$.

The somewhat similar, but more general structure for such systems is given as

$$Q_j \triangleq \bigvee_{Z_i \in \text{dom}(Z)} \left(\bigwedge_{k=1}^m (j_k(Z_i, H_j; D, S) \ni k_k(Z_i, H_j; D, S)) \right) \quad (5.1)$$

representing $(D \cdot S \ni H)$, where for all k, j_k and k_k are, possibly expert-derived, boolean functions, i.e., combinations of operations $\cdot, \vee, ()$.

Next, to complete the general data fusion theory again referring to Figure 1, we must choose an ALDP, i.e., a pair consisting of a compatible choice of formal language followed by a semantic evaluation or logic.

Consider then as reasonable candidates for the evaluation of (5.1), ALDP 1, 2, 3 as in Figure 1.

Again, it can be shown quite readily the first 3 ALDP examples in Figure 1 are such that their formal language components satisfy (a)-(d), Theorem 2, when implication is interpreted as

$$\ni \rightarrow \rightarrow, \quad (5.2)$$

where for all α, β

$$(\beta \rightarrow \alpha) \triangleq (\beta \vee \alpha). \quad (5.3)$$

Details of these evaluations are given in [1] Section 7. However, for fixed antecedents, it is seen there that negation and disjoint union (+) fail in all of these ALDP's to yield homomorphisms, but ALDP 4 (to be explained in the next section) does possess this property- indeed it is a characterizing relation.

Consider next:

$$\begin{aligned} p(\beta_0 \ni \alpha_0) &= p(\beta_0 \vee \alpha_0) = 1 - p((\beta_0' \vee \alpha_0')') = 1 - p(\beta_0' \cdot \alpha_0') \\ &= p(\alpha_0 | \beta_0) + p(\alpha_0' | \beta_0) - p(\beta_0' \cdot \alpha_0') \\ &= p(\alpha_0 | \beta_0) + p(\alpha_0' | \beta_0) - p(\alpha_0' | \beta_0) \cdot p(\beta_0') \\ &= p(\alpha_0 | \beta_0) + p(\alpha_0' | \beta_0) \cdot p(\beta_0') \\ &\geq p(\alpha_0 | \beta_0) \end{aligned} \quad (5.4)$$

$$\geq p(\alpha_0 \cdot \beta_0). \quad (5.5)$$

where the conditional probability is defined as usual as, e.g.,

$$p(\alpha_0 | \beta_0) \triangleq p(\alpha_0 \cdot \beta_0) / p(\beta_0), \quad (5.6)$$

provided $p(\beta_0) > 0$.

The above inequalities are strict, in general, and show that, basically, we cannot identify implication, as defined in the formal language (i) via eq. (5.2), with a "conditional object" such as $(\alpha_0 | \beta_0)$, otherwise this would, following evaluations by p and making the natural identification

$$p((\alpha_0 | \beta_0)) = p(\alpha_0 | \beta_0). \quad (5.7)$$

contradict the inequality in (5.4). Hence the behavior of conditional probabilities, while roughly resembling that of the probability of implications is not the same - indeed, one can, by choosing judiciously β_0 close to 0 in some natural sense, make $p(\beta_0 \Rightarrow \alpha_0)$ approach unity, while for the same choice of α_0, β_0 , $p(\alpha_0 | \beta_0)$ approaches zero. The significance of these results will be explored further in the next section, where we develop an ALDP (4) where formal implications $\alpha_0 \Rightarrow \beta_0$ can be identified with "conditional objects" $(\alpha_0 | \beta_0)$, whose semantic evaluations as in (5.7) are conditional probabilities; but in light of the above remarks, necessarily these entities lie outside of the original space of propositions Ω .

6. DATA FUSION AND CONDITIONAL OBJECTS

In Section 5, we have seen how a general inference rule structure for data fusion can be evaluated through three different approaches ALDP 1-3. In all of these, the key connector for inference \Rightarrow was interpreted in the formal language components as \rightarrow as given in eq. (5.3). On the other hand a natural - and commonly used - semantic evaluation for inference rules is through conditional probabilities. That is, the evaluation of a typical form $(j_{kij} \Rightarrow k_{kij})$ is $p(k_{kij} | j_{kij})$ for some choice of probability measure p over Ω , the set of all events or propositions, which for purposes of simplicity, from now on is assumed to be a boolean algebra. With this choice of evaluation, apropos to the spirit of this paper, we seek a formal language which will be compatible with these evaluations, i.e., will form an ALDP.

However, as pointed out in the discussion in the previous section centered around (5.4), one cannot identify implication via (5.2) with conditioning as evaluated in (5.7). The apparently commonly-held belief that such an identification can be made with no serious consequences, often called in the literature of logic as Stalnaker's Thesis [9], was attacked by Lewis [10] and independently by Calabrese [11]. The latter indeed showed, by use of a simple canonical expansion, that not only \rightarrow in (5.2) would not work, but any boolean function of two variables could not be used to play the role of conditioning, compatible with conditional probability evaluations.

Moreover, it would be particularly desirable, to replace this rather flawed situation, with an ALDP which would yield feasible computations for data fusion or at least be on the same order of complexity as ALDP 1,2,3. Note of course, if truly all inference rule antecedents are identical, as is the case essentially in Sections 3,4, then there is no real need to work with conditional objects, since all conditioned events can be simply considered as unconditional ones relative to their intersections with the fixed common antecedent, or one can stick with the interpretation of implication as in (5.2). (Compatible with this result, note the homomorphic relations for implication \rightarrow w.r.t. disjunction and conjunction - but not negation - as given in eqs. (7.4), (7.5) of [11].)

But, for the modeling of data fusion through inference rules with varying antecedents, no such direct

simplification occurs and the development of such conditional objects would address the problem. Although we have stated above that implication operator \rightarrow for a fixed antecedent yields homomorphic relations for \vee, \wedge , but not $()'$, conditional probabilities are compatible with homomorphic relations holding for all three operations, for any fixed antecedent; i.e., obviously, for all $\alpha_0, \beta_0, \gamma_0 \in \Omega$,

$$p((\alpha_0 | \gamma_0)') = 1 - p(\alpha_0 | \gamma_0) = p(\alpha_0' | \gamma_0), \quad (6.1)$$

$$p((\alpha_0 | \gamma_0) \vee (\beta_0 | \gamma_0)) = p(\alpha_0 \vee \beta_0 | \gamma_0), \quad (6.2)$$

$$p((\alpha_0 | \gamma_0) \cdot (\beta_0 | \gamma_0)) = p(\alpha_0 \cdot \beta_0 | \gamma_0). \quad (6.3)$$

Recall also the operation $+$ over Ω , which in terms of $\vee, \cdot, ()'$ is, for any $\alpha_0, \beta_0 \in \Omega$

$$\alpha_0 + \beta_0 = \alpha_0 \cdot \beta_0' \vee \alpha_0' \cdot \beta_0, \quad (6.4)$$

and conversely,

$$\alpha_0 \vee \beta_0 = \alpha_0 + \beta_0 + \alpha_0 \cdot \beta_0 \quad (6.5)$$

$$\alpha_0' = \alpha_0 + 1. \quad (6.6)$$

Noting that also, for any $\alpha_0, \beta_0 \in \Omega$,

$$p(\alpha_0 | \beta_0) = p(\alpha_0 \cdot \beta_0 | \beta_0), \quad (6.8)$$

the next result shows that under quite mild and simple conditions, conditional objects are essentially characterized:

Theorem 4. Characterization of conditional objects [12]

Given boolean ring Ω , there is a unique space $\bar{\Omega}$ of smallest possible classes - according to subset partial ordering - denoted as the conditional objects $(\alpha_0 | \gamma_0), (\beta_0 | \gamma_0), (\beta_0 | \zeta_0), \dots$, for all $\alpha_0, \beta_0, \gamma_0, \zeta_0, \dots \in \Omega$, such that the measure-free counterparts of (6.1)-(6.3) and (6.8) hold. That is,

$$(\alpha_0 | \gamma_0)' = (\alpha_0' | \gamma_0), \quad (6.9)$$

$$(\alpha_0 | \gamma_0) \vee (\beta_0 | \gamma_0) = (\alpha_0 \vee \beta_0 | \gamma_0), \quad (6.10)$$

$$(\alpha_0 | \gamma_0) \cdot (\beta_0 | \gamma_0) = (\alpha_0 \cdot \beta_0 | \gamma_0), \quad (6.11)$$

and equivalent to (6.9)-(6.11), one can require eqs. (6.11) and

$$(\alpha_0 | \gamma_0) + (\beta_0 | \gamma_0) = (\alpha_0 + \beta_0 | \gamma_0) \quad (6.12)$$

to hold; and

$$(\alpha_0 | \gamma_0) = (\alpha_0 \cdot \gamma_0 | \gamma_0). \quad (6.13)$$

Specifically, such conditional objects constitute all possible principal ideal cosets of ring Ω , where for any $\alpha_0, \gamma_0 \in \Omega$,

$$\begin{aligned} (\alpha_0 | \gamma_0) &= \alpha_0 \cdot \gamma_0' + \alpha_0 \\ &= \alpha_0 \cdot \gamma_0' + \alpha_0 \cdot \gamma_0 = \alpha_0 \cdot \gamma_0' \vee \alpha_0 \cdot \gamma_0 \\ &= \{x \cdot \gamma_0' + x_0 \cdot \gamma_0 | x \in \Omega\} \subseteq \Omega. \end{aligned} \quad (6.14)$$

the principal ideal coset generated by γ_0' with respect

due α_0 .

Proof: Use first the basic homomorphism theorem for quotient rings and the equivalence class property of cosets applied to (6.13). Again, see [12].

For a history of previous work in this area, see [1], Section 8.

In the approach taken here, developing all results from first principles considerations, the required operations upon conditional objects are defined simply as the natural class or component-wise extensions of the original operations. Thus, for example, let $\alpha_0, \beta_0, \gamma_0, \delta_0 \in \Omega$ arbitrary. The natural class extension of \cdot applied now to $(\alpha_0 | \beta_0) \cdot (\gamma_0 | \delta_0)$, noting each conditional object is in reality via (6.14) a subset of Ω , yields:

$$\begin{aligned} (\alpha_0 | \beta_0) \cdot (\gamma_0 | \delta_0) &= \{q \cdot r | q \in (\alpha_0 | \beta_0), r \in (\gamma_0 | \delta_0)\} \\ &= \{(x \cdot \beta'_0 + \alpha_0) \cdot (y \cdot \delta'_0 + \gamma_0) | x, y \in \Omega\} \\ &\subseteq \Omega. \end{aligned} \quad (6.15)$$

The basic structure of the conditional object extension $\bar{\Omega}$ of Ω is summarized next.

Theorem 5. Basic structure of $\bar{\Omega}$ [12], [13], [14].

(i) In terms of quotient rings,

$$\bar{\Omega} = \frac{u(\Omega/\Omega \cdot \gamma'_0)}{\gamma_0 \in \Omega} = \frac{u(\Omega/\Omega \cdot \gamma'_0)}{\gamma_0 \in \Omega}. \quad (6.16)$$

(ii) Conditioning as defined here can be identified essentially as the functional inverse of one-sided conjunction, i.e., conditional objects $(\alpha_0 | \gamma_0)$ all satisfy the basic relation analogous to (5.6) for conditional probabilities and a related condition:

$$(\alpha_0 | \gamma_0) \cdot \gamma_0 = \alpha_0 \cdot \gamma_0 \quad (6.17)$$

and

$$(\alpha_0 | \gamma_0) = \{x | x \in \Omega, x \cdot \gamma_0 = \alpha_0 \cdot \gamma_0\}. \quad (6.18)$$

(iii) The natural class extensions of all boolean operations from Ω to $\bar{\Omega}$ are well-defined/closed with ring-like properties, i.e., in the same previous sense, $\bar{\Omega}$ is a modified boolean algebra.

(iv) $\Omega \subseteq \bar{\Omega}$,

since for all $\alpha_0 \in \Omega$, (6.14) shows immediately that

$$(\alpha_0 | 1) = \{\alpha_0\}. \quad (6.19)$$

(v) Also, partial order \leq defined over $\bar{\Omega}$, characterized by, for any $\alpha_0, \beta_0 \in \bar{\Omega}$,

$$\alpha_0 \leq \beta_0 \text{ iff } \alpha_0 = \alpha_0 \cdot \beta_0 \text{ iff } \beta_0 = \beta_0 \vee \alpha_0, \quad (6.20)$$

can be extended directly to $\bar{\Omega}$ with the same characterizations as in (6.20) where (unconditional) objects in Ω are replaced by conditional ones in $\bar{\Omega}$. Then, combining this with (iii) and (iv) establishes $(\bar{\Omega}, \cdot, ()', +, \leq)$ as a natural extension of its unconditional counterpart $(\Omega, \cdot, ()', +, \leq)$.

(vi) A basic calculus of operations is, in addition to the properties in (6.9)-(6.13) for any $\alpha_i, \gamma_i \in \Omega$,

$i=1, \dots, m, m \geq 1$,

$$\bigvee_{i=1}^m (\alpha_i | \gamma_i) = \left(\bigvee_{i=1}^m \alpha_i \mid \bigvee_{i=1}^m \alpha_i \cdot \gamma_i \vee \bigvee_{i=1}^m \gamma_i \right), \quad (6.21)$$

$$\bigwedge_{i=1}^m (\alpha_i | \gamma_i) = \left(\bigwedge_{i=1}^m \alpha_i \mid \bigvee_{i=1}^m \alpha_i \cdot \gamma_i \vee \bigwedge_{i=1}^m \gamma_i \right), \quad (6.22)$$

$$\bigoplus_{i=1}^m (\alpha_i | \gamma_i) = \left(\bigoplus_{i=1}^m \alpha_i \mid \bigoplus_{i=1}^m \gamma_i \right). \quad (6.23)$$

Noting the reductions of (6.21)-(6.23) when antecedent $\gamma_1 = \dots = \gamma_m = \gamma_0$, as in (6.9)-(6.12), it follows that all boolean operational extensions over Ω coincide with corresponding coset operations when restricted to a fixed quotient ring, here $\Omega/\Omega \cdot \gamma'_0$.

(vii) As a special case of (6.22), the following chaining condition holds for all $\alpha_0, \beta_0, \gamma_0 \in \Omega$:

$$(\alpha_0 \cdot \beta_0 | \gamma_0) = (\beta_0 | \gamma_0) \cdot (\alpha_0 | \beta_0 \cdot \gamma_0). \quad (6.24)$$

Proof: The most difficult proof is that of (6.22). A sketch of the proof for the case $m=2$ is given in [13], Theorem 3.1, a full proof is presented in [12] where all other proofs are also given.

Remarks.

Apropos to Theorem 5(i), it follows that all results in the theory and application of linear (w.r.t. \cdot over \vee) boolean equations, such as presented in [15], can be reinterpreted in terms of conditional objects. Extensions of the concept of conditioning to more general structures than boolean, such as modified boolean, or Von Neumann regular, or to a category theory setting, have been considered [12].

Many other mathematical properties have been derived for conditional objects, including: characterizations for iterated conditional objects, i.e., conditional objects whose antecedent and consequent are also conditional objects; extensions of Stone's Representation Theorem to conditional objects; development of an outer approximation technique to force closure for non-boolean functions, including arithmetic operations over conditional objects; relations established between ordinary conditional random variables and a randomized version of conditional objects; and establishment of various probabilistic connections, such as measure-free independence; measure-free bayesian and sequential learning forms; and the proof that the extension of any probability measure $p: \Omega \rightarrow [0,1]$ to $p: \bar{\Omega} \rightarrow [0,1]$ through eq. (5.7) yields for the extension a monotone function. (Again, see [12]-[14], for further details.)

Most importantly here, analogues of calculus of relations for ALDP 1 ((1), (7.2)-(7.7)) hold for conditional objects, as Theorem 5 shows. Moreover, the hypotheses for Theorem 2 all hold here. At this point let us define ALDP 4, for a given boolean algebra Ω as simply

$$\text{ALDP 4} = (\bar{\Omega}, p), \quad (6.25)$$

where $p: \bar{\Omega} \rightarrow [0,1]$ is the conditional probability extension of $p: \Omega \rightarrow [0,1]$, as mentioned above and where

implication is interpreted as conditioning, i.e., for all $\alpha_0, \beta_0 \in \Omega$,

$$(\beta_0 \supset \alpha_0) = (\alpha_0 | \beta_0). \quad (6.26)$$

(Note that implication or conditioning here is restricted to be upon unconditional elements, i.e. elements of Ω , not upon other properly conditional objects. Some results indicate a possible identification of iterated conditional forms with simple conditional objects ([14], Section 4) so that in a sense this restriction may be unnecessary.)

Finally, consider use of ALDP 4 in evaluating data fusion expression Q in (5.1):

Direct use of (6.21) and (6.22) show that

$$\begin{aligned} p(Q=Q_j) &= p\left(\bigvee_{Z_j \in \text{dom}(Z)} \left(\bigwedge_{k=1}^m (k_{kij} | j_{kij}) \right)\right) \\ &= p(\Delta(H_j; D, S) | \Delta(H_j; D, S) \vee q(H_j; D, S)) \\ &= p(\Delta(H_j; D, S)) / p(\Delta(H_j; D, S) \vee q(H_j; D, S)), \end{aligned} \quad (6.27)$$

where

$$q(H_j; D, S) \triangleq \bigvee_{Z_j \in \text{dom}(Z)} \left(\bigwedge_{k=1}^m (k_{kij} \cdot j_{kij}) \vee \bigwedge_{k=1}^m j_{kij} \right) \quad (6.28)$$

and

$$\Delta(H_j; D, S) \triangleq \bigvee_{Z_j \in \text{dom}(Z)} \left(\bigwedge_{k=1}^m (k_{kij} \cdot j_{kij}) \right). \quad (6.29)$$

Thus, due to the calculus of operations given in Theorem 5, computations for data fusion using ALDP 4, with implication interpreted as a conditioning, compatible with conditional probabilities, appears no more complex than that for the other choices of ALDP's.

7. SUMMARY AND FUTURE DIRECTIONS

Because obviously C^3 systems are large scale ones, relatively few attempts have been made at approaching such systems from the viewpoint presented in this paper: a microscopic bottom's up approach. Indeed, the system branching problem is so formidable, that for any realistic implementation, pruning techniques are necessary at all stages of modeling. However, this should not preclude anyone - if even naively - from attempting a first cut, first-principles approach, without any such abridgment. Such a theoretical structure can serve as a common framework or language for comparing and contrasting entire or parts of C^3 "theories". Furthermore, this generic model could be useful in rigorizing overall C^3 activities within a unified framework and provide insights not available from a situation-specific or only local level. In a similar vein, data fusion is considered here as an integral part of a larger C^3 structure and identified with the combination of evidence problem.

As for future directions, further work must be done in integrating the cognitive process phase with the full semantic evaluation carried out for choice of an ALDP. This should include mental imaging and related thought processes. Alternative data fusion structures, such as recursive forms analogous to Kalman filter forms, will also be considered. Tie-ins with proposed macroscopic C^3 models have yet to be carried out.

8. ACKNOWLEDGMENTS

This work was supported jointly by the Naval Ocean Systems Center Program for Independent Research (IR) and the Joint Directors of Laboratories Technical Panel for C^3 , Basic Research Group (JDL,TPC³,BRG).

9. REFERENCES

1. I.R. Goodman, "A general theory for the fusion of data", *Proceedings First Tri-Service Data Fusion Symp. (J.D.L.)*, held at Applied Physics Lab., Laurel, Md., June 9-11, 1987.
2. M. Kaku and J. Trainer, *Beyond Einstein*, Bantam Books, New York, March, 1987.
3. I.R. Goodman, *PACT: An Approach to Combining Linguistic-Based and Probabilistic Information for Correlation and Tracking*, NOSC Tech. Doc. 878, March, 1986, Naval Ocean Systems Center, San Diego, CA.
4. I.R. Goodman, "A probabilistic / possibilistic approach to modeling C^3 systems", *Proc. 9th MIT/ONR Workshop. C³ Sys.*, Dec., 1986, 53-58.
5. I.R. Goodman and H.T. Nguyen, *Uncertainty Models for Knowledge-Based Systems*, North-Holland Co., Amsterdam, 1985.
6. C.R. Rao, *Linear Statistical Inference and Its Applications*, 2nd ed., John Wiley Co., New York, 1973.
7. G. Shafer, *A Mathematical Theory of Evidence*, Princeton Univ. Press, 1976.
8. H.T. Nguyen, "On random sets and belief functions", *J. Math. Analysis & Applic.*, 65, 1978, 531-542.
9. R.C. Stalnaker, "A theory of conditionals", *Amer. Phil. Quarterly, Monograph Series No.2*, N. Rescher ed., Oxford, 1968, 98-112.
10. D. Lewis, "Probabilities of conditionals and conditional probabilities", *The Phil. Rev.*, LXXIV, No.3, July, 1976, 297-315.
11. P. Calabrese, "An algebraic synthesis of the foundations of logic and probability", to appear in *Information Sciences*, 1987-88.
12. I.R. Goodman and H.T. Nguyen, *An Algebraic Theory of Conditioning with Applications to Uncertainty Modeling*, monograph to be submitted, 1987-88.
13. I.R. Goodman, "A measure-free approach to conditioning", *Proc. 3rd AAAI Workshop Uncert. A.I.*, 1987, based on conference held at Seattle, July 10-12, 1987.
14. I.R. Goodman and H.T. Nguyen, "Conditional objects and the modeling of uncertainties", to appear in *Fuzzy Computing*, M.M. Gupta et al., eds., North-Holland Co., New York, 1987-88.
15. S. Rudeanu, *Boolean Functions and Equations*, North-Holland Co., Amsterdam, 1974.

END

DATE

FILMED

5-88
DTIC